



[About](#) [Citation Policy](#) [Donate a Data Set](#)
[Contact](#)

☒ Repository ☐ Web 

[View ALL Data Sets](#)

Dermatology Data Set

Download: [Data Folder](#), [Data Set Description](#)

Abstract: Aim for this dataset is to determine the type of Eryhemato-Squamous Disease.

Data Set Characteristics:	Multivariate	Number of Instances:	366	Area:	Life
Attribute Characteristics:	Categorical, Integer	Number of Attributes:	33	Date Donated	1998-01-01
Associated Tasks:	Classification	Missing Values?	Yes	Number of Web Hits:	71633

Source:

Original Owners:

1. Nilsel Ilter, M.D., Ph.D.,
 Gazi University,
 School of Medicine
 06510 Ankara, Turkey
 Phone: +90 (312) 214 1080

2. H. Altay Guvenir, PhD.,
 Bilkent University,
 Department of Computer Engineering and Information Science,
 06533 Ankara, Turkey
 Phone: +90 (312) 266 4133
 Email: guvenir '@' cs.bilkent.edu.tr

Donor:

H. Altay Guvenir,
 Bilkent University,
 Department of Computer Engineering and Information Science,
 06533 Ankara, Turkey
 Phone: +90 (312) 266 4133
 Email: guvenir '@' cs.bilkent.edu.tr

Data Set Information:

This database contains 34 attributes, 33 of which are linear valued and one of them is nominal.

The differential diagnosis of erythematous-squamous diseases is a real problem in dermatology. They all share the clinical features of erythema and scaling, with very little differences. The diseases in this group are psoriasis, seborrheic dermatitis, lichen planus, pityriasis rosea, chronic dermatitis, and pityriasis rubra pilaris. Usually a biopsy is necessary for the diagnosis but unfortunately these diseases share many histopathological features as well. Another difficulty for the differential diagnosis is that a disease may show the features of another disease at the beginning stage and may have the characteristic features at the following stages. Patients were first evaluated clinically with 12 features. Afterwards, skin samples were taken for the evaluation of 22 histopathological features. The values of the histopathological features are determined by an analysis of the samples under a microscope.

In the dataset constructed for this domain, the family history feature has the value 1 if any of these diseases has been observed in the family, and 0 otherwise. The age feature simply represents the age of the patient. Every other feature (clinical and histopathological) was given a degree in the range of 0 to 3. Here, 0 indicates that the feature was not present, 3 indicates the largest amount possible, and 1, 2 indicate the relative intermediate values.

The names and id numbers of the patients were recently removed from the database.

Attribute Information:

Clinical Attributes: (take values 0, 1, 2, 3, unless otherwise indicated)

- 1: erythema
- 2: scaling
- 3: definite borders
- 4: itching
- 5: koebner phenomenon
- 6: polygonal papules
- 7: follicular papules
- 8: oral mucosal involvement
- 9: knee and elbow involvement
- 10: scalp involvement
- 11: family history, (0 or 1)
- 34: Age (linear)

Histopathological Attributes: (take values 0, 1, 2, 3)

- 12: melanin incontinence
- 13: eosinophils in the infiltrate
- 14: PNL infiltrate
- 15: fibrosis of the papillary dermis
- 16: exocytosis
- 17: acanthosis
- 18: hyperkeratosis
- 19: parakeratosis
- 20: clubbing of the rete ridges
- 21: elongation of the rete ridges
- 22: thinning of the suprapapillary epidermis
- 23: spongiform pustule
- 24: munro microabscess
- 25: focal hypergranulosis
- 26: disappearance of the granular layer
- 27: vacuolisation and damage of basal layer
- 28: spongiosis
- 29: saw-tooth appearance of rete
- 30: follicular horn plug
- 31: perifollicular parakeratosis
- 32: inflammatory mononuclear infiltrate
- 33: band-like infiltrate

Relevant Papers:

G. Demiroz, H. A. Govenir, and N. Ilter, "Learning Differential Diagnosis of Eryhemato-Squamous Diseases using Voting Feature Intervals", Artificial Intelligence in Medicine
[\[Web Link\]](#)

Papers That Cite This Data Set¹:



Vassilis Athitsos and Stan Sclaroff. [Boosting Nearest Neighbor Classifiers for Multiclass Recognition](#). Boston University Computer Science Tech. Report No, 2004-006. 2004. [\[View Context\]](#).

Gisele L. Pappa and Alex Alves Freitas and Celso A A Kaestner. [Attribute Selection with a Multi-objective Genetic Algorithm](#). SBIA. 2002. [\[View Context\]](#).

Rafael S. Parpinelli and Heitor S. Lopes and Alex Alves Freitas. [An Ant Colony Based System for Data Mining: Applications to Medical Data](#). CEFET-PR, CPGEI Av. Sete de Setembro, 3165. [\[View Context\]](#).

Gisele L. Pappa and Alex Alves Freitas and Celso A A Kaestner. [AMultiobjective Genetic Algorithm for Attribute Selection](#). Computing Laboratory Pontificia Universidade Catolica do Parana University of Kent at Canterbury. [\[View Context\]](#).

Perry Moerland. [Mixtures of latent variable models for density estimation and classification](#). E S E A R C H R E P
R O R T I D I A P D a l l e M o l l e I n s t i t u t e f o r P e r c e p t u a l A r t i f i c i a l I n t e l l i g e n c e . [\[View Context\]](#).

H. Altay Guvenir. [A Classification Learning Algorithm Robust to Irrelevant Features](#). Bilkent University, Department of Computer Engineering and Information Science. [\[View Context\]](#).

M. V. Fidelis and Heitor S. Lopes and Alex Alves Freitas. [Discovering Comprehensible Classification Rules with a Genetic Algorithm](#). UEPG, CPD CEFET-PR, CPGEI PUC-PR, PPGIA Praa Santos Andrade, s/n Av. Sete de Setembro. [\[View Context\]](#).

Rafael S. Parpinelli and Heitor S. Lopes and Alex Alves Freitas. [PART FOUR: ANT COLONY OPTIMIZATION AND IMMUNE SYSTEMS Chapter X An Ant Colony Algorithm for Classification Rule Discovery](#). CEFET-PR, Curitiba. [\[View Context\]](#).

Citation Request:

Please refer to the Machine Learning Repository's [citation policy](#)

[1] Papers were automatically harvested and associated with this data set, in collaboration with [Rexa.info](#)

Supported By:



In Collaboration With:



[About](#) || [Citation Policy](#) || [Donation Policy](#) || [Contact](#) || [CML](#)